Logistic Regression: Equations

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

Equations

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N},$$

$$\text{where } x_{1:N} \in \mathbb{R}^{N \times n} \text{ and } y_{1:N} \in \mathbb{R}^{N \times m}.$$

$$\hat{y}_{1:N} = \sigma(x_{1:N} \cdot W + b)$$

$$\hat{P}(\mathrm{y} = \mathrm{True}|x_i) \approx \hat{y}_i$$

$$\mathcal{L}(W, b) = -\frac{1}{N} \sum_{i=1}^{N} y_i^\top \cdot \log \hat{y}_i + (1 - y_i)^\top \cdot \log(1 - \hat{y}_i)$$

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \mathcal{L}(\theta), \text{ where } \theta = \{W, b\}.$$

$$W \leftarrow W - \eta \nabla_W \mathcal{L}(W, b)$$

$$b \leftarrow b - \eta \nabla_b \mathcal{L}(W, b)$$

Equations

$$\begin{aligned}
\text{BCELoss}(y_{1:N}, \hat{y}_{1:N}) &= -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{m} y_{i,j} \times \log \hat{y}_{i,j} + (1 - y_{i,j}) \times \log\left(1 - \hat{y}_{i,j}\right) \\
&= -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{m} P(\mathrm{y}_j = \text{True}|x_i) \times \log P(\mathrm{y}_j = \text{True}|x_i;\theta) + P(\mathrm{y}_j = \text{False}|x_i) \times \log P(\mathrm{y}_j = \text{False}|x_i;\theta), \\
&\qquad \text{where } y_{i,j} = P(\mathrm{y}_j = \text{True}|x_i) \text{ and } 1 - y_{i,j} = P(\mathrm{y}_j = \text{False}|x_i).
\end{aligned}$$